

DOCUMENT RESUME

ED 080 968

CS 000 706

AUTHOR Hintzman, Douglas L.
TITLE Memory Judgments. Final Report.
INSTITUTION Oregon Univ., Eugene. Dept . of Psychology.
SPONS AGENCY National Center for Educational Research and
Development (DHEW/OE), Washington, D.C.
BUREAU NO BR-1-0343
PUB DATE Aug 73
GRANT OEG-X-71-0033 (508)
NOTE 52p.
EDRS PRICE MF-\$0.65 HC-\$3.29
DESCRIPTORS *Cognitive Processes; Educational Research; *Memory;
Perception; Reading; *Recall (Psychological);
*Recognition; *Retention

ABSTRACT

The usefulness of the method of memory judgments as a tool for studying human memory was explored in this research. The first experiments involved the nature of the information about when an event occurred, the spacing of repetitions of an event, and presentation modality of an event. Results suggested that time and spacing are represented in memory by contextual associations, while modality is represented, in addition, by a literal copy of the perceptual experience. A second group of experiments involved the effect of the spacing of repetitions on retention. They suggested that the locus of the effect is on the second of two presentations of the event and that the effect is primarily a function of duration of the spacing interval. The third group of experiments involved memory for the duration of an interval. Judgments of duration were not found to be related in any simple way to standard recall or recognition measures of memory for events during the interval. (Author/LL)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

CS

Final Report

Project No. 1-0343
Grant No. OEG-X-71-0033(508)

Douglas L. Hintzman
Department of Psychology
University of Oregon
Eugene, Oregon 97403

MEMORY JUDGMENTS

August 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

ED 080963

CS0000706

Final Report

Project No. 1-0343
Grant No. OEG-X-71-0033(508)

Douglas L. Hintzman
Department of Psychology
University of Oregon
Eugene, Oregon 97403

MEMORY JUDGMENTS

August 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Center for Educational Research and Development

Author's Abstract

This grant supported 17 experiments, most of which explored the usefulness of the method of memory judgments as a tool for studying human memory. The first experiments were concerned with the general question, "What is stored?" Specifically, they involved the nature of the information about when an event occurred, the spacing of repetitions of an event, and the presentation modality of an event. Results suggest that time and spacing are represented in memory by contextual associations, while modality is represented, in addition, by a "literal copy" of the perceptual experience. A second group of experiments involved the effect of the spacing of repetitions on retention. They suggest that the locus of the effect is on the second of two presentations of the event, and that the effect is primarily a function of duration of the spacing interval. Experiments aimed at supporting explanations of the effect in terms of rehearsal, voluntary control of attention, and habituation and recovery, all had negative outcomes. A third topic involved memory for the duration of an interval. Judgments of duration were not found to be related in any simple way to standard recall or recognition measures of memory for events during the interval.

Final Report

Project No. 1-0343
Grant No. OEG-X-71-033(508)

MEMORY JUDGMENTS

Douglas L. Hintzman
University of Oregon
Eugene, Oregon

August 1973

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
National Center for Educational Research and Development

Table of Contents

List of Tables - - - - -	5
List of Figures - - - - -	6
Introduction - - - - -	7
Method and Results - - - - -	10
Conclusions - - - - -	45
Bibliography - - - - -	49

List of Tables

Table		Page
1	Mean of Median Correct Reaction Times and Error Percentages, Experiment IV.	19
2	Response Proportions and Mean Judgments of Frequency, Experiment V.	21
3	Mean Confidence Ratings and Response Percentages, Experiment XV.	42

List of Figures

Figure		Page
1	Corrected response probabilities for the 12 conditions of Experiment I.	12
2	Median spacing judgments, Experiment II.	15
3	Mean judgments of frequency as a function of spacing and modality condition, Experiment VI.	24
4	Component analysis of judgments for Conditions AV and VA; Experiment VI.	26
5	Mean judged frequency of pictures as a function of frequency and the spacing of repetitions; Experiment VII.	29
6	Combined data of Experiments VII, VIII, and IX: Mean judged frequency as a function of duration of the spacing interval.	31
7	Effects of spacing and attention given P_2 on judged frequency of pictures; Experiment XII.	36
8	Effects of spacing and P_1 duration on judged frequency of pictures; Experiment XIII.	38
9	Effects of number of prior massed presentations and spacing on judged frequency; Experiment XIV.	40

Introduction

Advances in scientific understanding very often are triggered by improvements in the technical and methodological tools available to the investigator. The recent upsurge of interest in the experimental study of human memory, in large part, results from a proliferation of experimental techniques. Recognition memory, free recall, verbal discrimination learning, and various short-term retention tasks, virtually ignored for many years, are today radically altering our conception of the nature of memory. It can be argued that, historically, the study of memory has suffered less from a dearth of relevant theoretical questions than it has from a limited stock of methodological techniques used to answer the questions.

The research reported here involves a deliberate attempt to expand the stock of methodological tools used in the investigation of memory. While the experiments that will be reported were directed at several different theoretical questions that are only loosely related, most of them made use of an experimental technique--actually a set of techniques--which we call collectively the "method of memory judgments". Not all examples of the method originated with the present investigator; judgments of recency, for example, were used by Yntema and Trask (1963) and later by Peterson (1967), discrimination of list membership has been employed by Winograd (1968) and others, and even the time-honored recognition memory task requires an explicit judgment (old or new) on the part of the subject, and so could be considered a simple example of the method of memory judgments. Nevertheless, the present approach is unique in recognizing the common elements in these methods and in developing new judgment tasks, whenever appropriate, which fall within the same general methodological framework. A primary motivation underlying the present project is the hope that these methods will be of value in investigating questions other than the particular ones under consideration here.

The method of memory judgments is essentially a technique for determining what information is stored in memory. The traditional method of determining "What is stored?" employs a transfer design (Postman, 1972). The subject learns Task A, then learns Task B, which bears some specific relationship with A. Comparison with a control condition presumably reveals whether the task components shared by A and B were stored during the learning of A. The transfer design can be used with animal subjects and with human subjects in tasks having little cognitive involvement, as well as in cognitive memory research. A more recent innovation, used almost exclusively in memory research, involves the study of errors committed by the subject. Relationships between the to-be-remembered item and the incorrectly-given item presumably can be used to infer the nature of the subject's memory code (e.g., acoustic, visual, or semantic). In recall tasks, spontaneously committed errors may be the basis of the inference (e.g., Conrad, 1964); in other tasks, such as recognition memory, the experimenter may attempt to force certain types of errors that he feels may be particularly revealing, by using as distractor items

stimuli having a specific relationship to those the subject was asked to remember (e.g., Anisfeld & Knapp, 1968). More recently still, reaction times have been used to determine the nature of the stored information (e.g., Posner, Boies, Eichelman, & Taylor, 1969). Recognition of a stimulus as old is thought to be faster if the test stimulus bears a direct relationship to the format of storage than if it does not. In the latter case, correct recognition depends upon certain time-consuming transformations, either of the input or of the stored information.

In contrast to these techniques, the method of memory judgments appears deceptively simple. The experimenter, essentially, asks the subject to describe in a very restricted way the information stored in memory. Usually the response is a numerical judgment (e.g., what position in the list you just saw was occupied by this word?) or a categorical judgment (e.g., was this word presented auditorily or visually?). Of course, the method does not really provide the direct access to the format of storage that this description suggests. The experimenter must choose the specific question that the subject is to be asked; and inferences still must be drawn from the subject's response concerning the nature of the stored information that gave rise to the judgment. Still, the method can be adapted to a variety of questions that would be difficult to answer using other techniques. And it is not intended to replace other methods, but rather to be used in conjunction with them. Hopefully, the more different ways we have of studying memory, the more complete picture, and therefore the more adequate theory, we will eventually obtain.

The experiments conducted during this project fall into three general categories. The first category consists of investigations of "What is stored?" Experiment I tested the hypothesis that the "time tag"--that stored information that tells a person approximately when a particular remembered event occurred--consists of associations between the event and the cognitive context that prevailed at the time of the event. Experiment II concerned a phenomenon previously discovered in our laboratory (Hintzman & Block, 1973)--the fact that subjects can remember with some accuracy the spacing of two repetitions of the same word but are unable to judge equivalent spacings of single presentations of two different words. The experiment tested a particular hypothesis regarding how the spacing information is stored. Experiments III and IV had to do with memory for mode of input. In one, it was asked whether subjects could remember which voice--male or female--had spoken a given word during the study phase. In the other, recognition latency was used to determine the extent to which the memory format for visually-presented words includes information about the type face in which the words were presented. Experiment V asked whether subjects could simultaneously remember two input modalities of a repeated word. The question is related to the hypothesis that repetition results in multiple traces, rather than a strengthening of traces (Hintzman & Block, 1971).

The second category consists of experiments investigating a particular phenomenon known as the spacing effect. Briefly, the effect is this: if two presentations of a to-be-remembered item (P_1 and P_2) occur close together in time, the item is more poorly remembered than it is if the two occurrences are spaced further apart. Several hypotheses concerning the cause of the spacing effect were reviewed by the present investigator in a paper read at the Second Loyola Conference on Cognition, in Chicago, in May 1973. The ten experiments (Experiments VI through XV) on the spacing effect presented here were designed to test a number of these hypotheses. While no definitive answer can be given yet regarding the cause of the spacing effect, these investigations appear to have narrowed somewhat the set of possible interpretations.

The third category consists of two experiments (XVI and XVII) done by Richard A. Block, as a doctoral dissertation. The experiments have to do with the experience of duration in retrospect--that is, memory for how long an interval was--and its relationship to performance on standard measures of retention of events in the interval.

METHOD AND RESULTS

Experiment I: Effects of delay on memory for serial position.

This study was done in conjunction with Richard A. Block and Jeffery J. Summers, and is Experiment III in an article appearing in the Journal of Experimental Psychology, 1973, 97, 220-229.

The question under consideration here concerned the nature of the "time tag"--the stored information that tells the subject when a particular remembered event occurred. One possibility is that temporal information is inferred from retrieved contextual information. For example, Anderson and Bower (1972) have assumed that when a word is presented it becomes associated with elements of the cognitive context in which it is embedded. Such associations, when retrieved on a later retention test, may provide information about the position the word occupied in the list.

This contextual association hypothesis makes at least one prediction that is different from those made by other explanations of the time tag. If several lists are presented, and the subject is then presented with test words and asked to indicate, for each, both the list in which it occurred and the position it occupied within the list, responses should not show a strict temporal generalization gradient. That is, to the extent that contextual elements can be manipulated independently of time, similar contextual elements can be expected to prevail at corresponding serial positions in different lists, so that words assigned to incorrect lists should tend to be assigned to the correct position, rather than the position temporally closest to the actual point of presentation.

A previous experiment, (Hintzman, Block, & Summers, 1973; Experiment II) provided evidence favoring the contextual association hypothesis except in one detail. Data showed that words misplaced in the incorrect list were indeed placed in the appropriate parts of the list. In addition, position judgments were better for words from the beginning of List 1 than for words from the middle of the series, and this is consistent with the notion that cognitive context changes most rapidly during the first few minutes of the experiment and more slowly thereafter. However, judgments for words from the end of the final list were as good as those from the beginning of List 1, a fact seemingly inconsistent with the hypothesis. The strong "recency effect" revealed in the position judgments from the end of the final list suggested an analogy with free recall. In free recall, a strong recency effect is found, but if a short delay is introduced before recall is taken, the recency effect disappears (Glanzer & Cunitz, 1966; Postman & Phillips, 1965). With this analogy with free recall in view, an experiment was run in which half the subjects were given the position-judgment test immediately after List 3, while the other half was not given the test until a delay of 15 min had ensued. The purpose was to study effects of the delay on the primacy, temporal, within-list position, and recency factors in position judgments.

The subjects, 90 paid volunteers tested in groups of up to 7 subjects, were presented with three lists of 27 words each. Following Lists 1 and 2, recognition tests were given in which the 13 words from the middle of the preceding list were mixed in with 13 new words, and the subject was to indicate which words were old. Following List 3, the treatment of the two groups of subjects differed. The Immediate group was immediately given the position judgment test, while the Delayed group was given a recognition test on the middle words from List 3, and then worked on an unrelated filler task for a total delay of approximately 15 min before being given the position judgment test. On the position judgment test, subjects were shown 50 words in random order. Eight of the words were new, and 7 each were from the beginnings and ends of the three experimental lists. Subjects were to respond for each word, either: New, 1B, 1E, 2B, 2E, 3B, or 3E, with the digit indicating the list number and B and E indicating beginning and end, respectively. The judgments, corrected for response bias by using a posteriori probabilities, are presented in Figure 1.

Results of the Immediate test are in the top panels of the figure; those of the Delayed test are in the bottom panels. For either group, the distribution of judgments for each of the six presentation conditions (3 lists x 2 positions) is presented. Data for the Immediate test show: (a) a primacy effect, indicated by the high accuracy in condition 1B; (b) a temporal factor, revealed by the generalization gradients around the correct response values in all three panels; (c) a within-list context effect, manifested in the tendency for the dashed curves to have more positive slopes than the solid curves, which indicates that words judged to have occurred in the wrong list nevertheless tended to be placed in the correct within-list position (beginning vs. end); and (d) a recency effect, seen in the high accuracy of position judgments for condition 3E. Data of the Delayed group are nearly identical with those of the Immediate group, except for (d), the recency effect. As suspected, the delay largely eliminated the recency effect, as it does in free recall.

The results of this experiment support the following interpretation: The primacy effect, generalization along the temporal dimension, and the within-list context effect may all be mediated by essentially the same mechanism, since none of these factors was affected by delay. We may hypothesize that contextual associations play the mediating role in these effects, with different elements entering into the within-list position effect than enter into the primacy and temporal generalization effects. The recency effect is apparently mediated by a different mechanism--possibly assessment of memory strength, although there are other possibilities (see discussion by Hintzman, Block, & Summers, 1973).

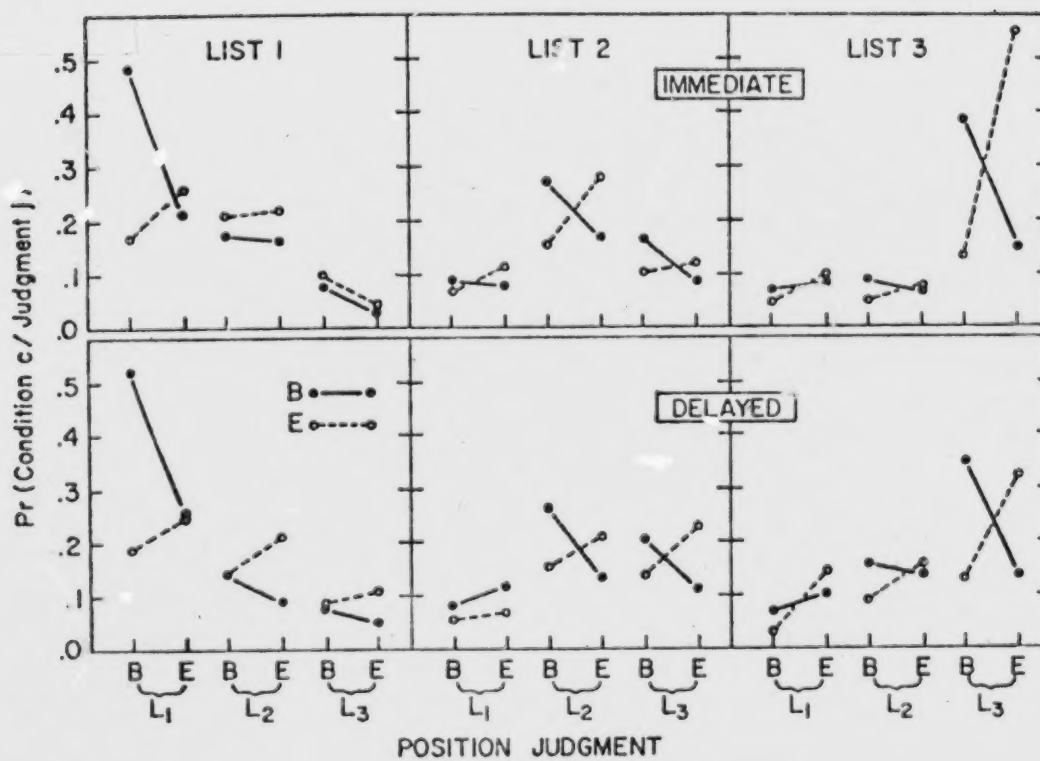


Figure 1. Corrected response probabilities for the 12 conditions (3 lists x 2 positions x 2 delays) of Experiment I.

Experiment II: Memory for the spacing of related words:

In a previous investigation (Hintzman & Block, 1973), a word list was presented in which a number of words occurred two times, with the spacing of repetitions varied. Matched with each repeated word was a word pair. Each member of the pair occurred once, and the spacing between the first and second members was the same as that of the repeated word with which the pair was matched. On the retention test, subjects were asked to judge the number of items that had intervened between two events--either instances of the same word (e.g., ELK-ELK) or of two different words (e.g., FAN-COT). The results showed that the subjects would judge the spacing of same-word pairs with some accuracy, but were unable to judge the spacing of different words. The explanation of this result offered by Hintzman and Block (1973) was as follows: when the second presentation of a word (P_2) occurs, it typically retrieves the trace of the first presentation (P_1). Thus, at the time of P_2 , the subject is aware that the word occurred before, and knows something about how long ago it occurred, based on the retrieved time tag. This knowledge could be characterized as an implicit judgment of recency, which constitutes part of the context of P_2 , and is therefore stored as part of the trace of P_2 . On the later judgment test, presentation of the word retrieves the traces of P_1 and P_2 ; and the trace of P_2 includes the implicit recency judgment, which constitutes information regarding the P_1 - P_2 spacing. Since this is the information requested of the subject, he is able to comply. In the Different Word condition, of course, the second word does not ordinarily retrieve the trace of the first, and so the same information is not stored.

This reasoning leads to an obvious prediction, which the present experiment was designed to test. If the two different words which constitute a test pair are related words, so that one will retrieve the trace of the other, then the recency judgment upon which the later spacing judgment is based should occur when the second member of the pair is presented. Thus, subjects should be able to judge the spacing of word pairs that are associatively related.

In this experiment, there were three kinds of test pairs, matched for spacing in the input sequence: Same Words, Related Words, and Unrelated Words. A pool of 96 nouns was generated such that there were 48 pairs of highly associated items. Examples are: WAR-PEACE, TABLE-CHAIR, KING-QUEEN, SPIDER-WEB. In the presentation list, words were selected from this pool and four pairs were assigned to each of 12 different conditions which combined the three types of test pair (Same, Related, and Unrelated) with four spacings: $S = 0, 3, 10$, and 25 intervening items. On the retention test form, word pairs of the three different types appeared, and the subjects made judgments of the number of words that intervened between the two members of the pair. Example pairs are: Same, TABLE-TABLE, Related, SPIDER-WEB, Unrelated, PEACE-KING. The subjects were 55 paid volunteers.

The median spacing judgments for all 12 conditions are presented in Figure 2. The results of Hintzman and Block (1973) were replicated, as is seen in the comparison of the Same Word and Unrelated Word curves. Again, spacing judgments for Same Word pairs did not vary with spacing. The condition added in this study, the Related Words condition, produced precisely the effect predicted by the hypothesis. Subjects were able to judge spacings of Related Word pairs--and in fact, they appear to have judged them as accurately as Same Word pairs. Presumably, this is because in a pair like SPIDER-WEB, the second member of the pair (WEB), when it occurs, retrieves the trace of the first (SPIDER), so that the subject is aware when storing WEB of the recency of SPIDER. The implicit recency judgment is the basis for the spacing judgment on the later test.

The implications of this finding should be mentioned. One is that memory theorists should be careful not to confuse the Study and Test phases of an experiment with the hypothetical encoding and retrieval processes. If our interpretation of the outcome is correct, subjects are continually retrieving information during the study phase of the experiment, and a great deal of the information they have in storage at the time of the test must be attributed to combined operation of encoding and retrieval processes rather than encoding processes alone. The second implication regards application. If spacing judgments do in fact indicate unintentional study-phase retrieval, they may prove to be of some value in investigating the organizational processes that go on when a subject studies a categorized word list and that evidence themselves in clustering in free recall. We hope to use spacing judgments to study organization processes in future research.

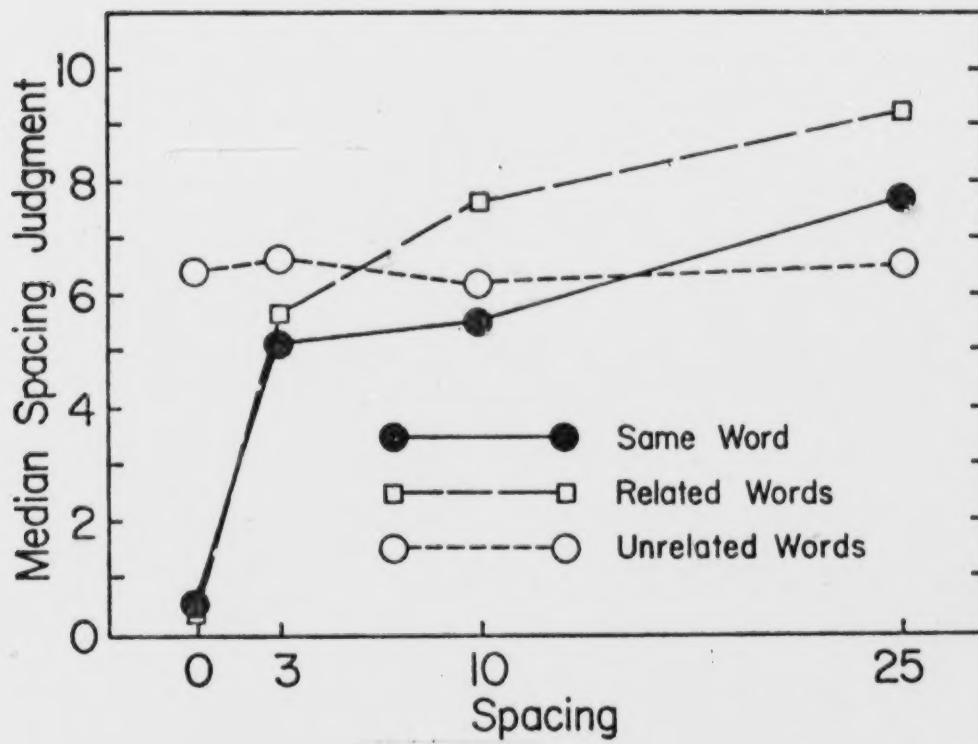


Figure 2. Median spacing judgments, Experiment II.

Experiment III: Memory for voicing.

This study was done in conjunction with Richard A. Block and Norman R. Inskeep, and is Experiment III in an article appearing in the Journal of Verbal Learning and Verbal Behavior, 1972, 11, 741-749.

This experiment and the next were concerned with subjects' memory for information about the mode of input of a word in a situation in which the subjects did not know they would be asked to remember this information. In a previous experiment (Hintzman, Block, & Inskeep, 1972; Experiment I) it was shown that subjects remember fairly well the input modality (auditory vs. visual) of a given word. This was demonstrated both by an index of clustering according to input modality in free recall and by a later modality-judgment test given on words from several previous free-recall lists. In a second experiment using purely visual presentation (Hintzman, Block, & Inskeep, 1972; Experiment II) it was found that subjects also remembered the type style in which visual words had been presented. Both clustering scores in free recall and judgments of type style indicated this. However, memory for type style, according to both measures, was not as good as memory for the auditory-visual distinction. The present experiment was designed to extend the basic finding to purely auditory presentation. Half the words were presented by a male voice (M) and half by a female voice (F), and subjects were later asked to judge which voice had presented a given word.

Subjects were 34 paid volunteers (17 males and 17 females). They were presented with eight lists, pre-recorded on tape, each 18 words long. In each list, half the words, randomly chosen, were read by a male voice and half were read by a female voice. Presentation was at a 3-sec rate. Following presentation of each list, subjects were given 3 min for free recall of the words in that list. Finally, a judgment test was administered, consisting of a test on which 160 words appeared --the 144 words from the eight lists plus 16 new words. For each word on the test, subjects were to choose one of three categories: New, Male, or Female. For half the subjects, the M and F voices were presented over different speakers located on opposite sides of the room. Thus location was a redundant cue with voicing. For the other half, both voices were heard at the same location.

Clustering in free recall was measured using the clustering index of Hudson and Dunn (1969) and also using a conditional probability analysis that is free of contamination by output priority effects. Since both measures led to the same conclusion, only the Hudson-Dunn measure will be presented here (for additional details, see Hintzman, Block, & Inskeep, 1972). The mean clustering index (where .00 indicates no clustering) was .15. This value differed significantly from .00 by a *t*-test, indicating organization of free recall by input voice. By contrast, clustering values for the two previous studies were .19 for auditory vs. visual presentation, and .10 for block vs. script type in the purely-visual presentation experiment. Apparently, within-

modality features do determine recall organization, but apparently not as strongly as the modality distinction does.

The percentage of correct identifications of input voice, given that the word was recognized as old, was 59%, where chance would be 50%. This compares with 74% accuracy on auditory vs. visual modality judgments and 58% on block vs. script typestyle judgments, in the previous experiments. Again, while the within-modality distinctions are retained, they are not remembered as well as the between-modality distinction. In this study, it did not matter whether the male and female voices had come from different or from the same location, and so the redundant location cue was either not stored or was not useful in the judgment task. Although female subjects were slightly better at remembering input voice, the sex of the subject did not interact with input voice as one might expect if rehearsal involved imagining the word occurring in one's own voice.

This experiment, like the auditory-visual and type-style experiments using the same design, shows that perceptual features of words are retained far longer than most theories have assumed. Apparently long-term memory for a word consists of more than semantic information. Information about the perceptual event itself is also retained, and can serve as the basis for a judgment of mode of input even after a retention interval of several minutes duration. The unanswered question is just how this perceptual information is represented, and that was the subject of the next experiment.

Experiment IV: Long-term visual traces of visually presented words.

This study was done in conjunction with Jeffery J. Summers, and appears in the Bulletin of the Psychonomic Society, 1973, 1, 325-327.

There are at least two ways in which information about mode of input (e.g., modality, type style, and voice) could be represented in memory. One is that mode of input, when noticed by the subject, is encoded as an abstract proposition associated with the meaning of the word, bearing only a referential relationship to the actual stimulus conditions. The other is that a literal copy of the perceptual experience persists, in visual or auditory memory, for several minutes after presentation, and thus outlives what is ordinarily considered to be sensory or even short-term memory.

One way to distinguish between these two hypotheses is to examine the joint effects of input and test modes on recognition. If the abstract proposition hypothesis is true, it should make no difference whether the original-presentation and test-presentation modes are the same. If the literal copy hypothesis is true it might, since a physical match between the current stimulus and the stored trace could facilitate recognition. The purpose of this experiment was to test the literal copy hypothesis by providing evidence regarding the role of physical identity of presentation and test stimuli on recognition memory. Words were presented visually, and recognition latencies were measured to determine whether the recognition decision is facilitated by physical identity.

The subjects were 52 paid volunteers. They were tested individually, with the experiment controlled by a computer and stimuli appearing on an oscilloscope display. During the study phase of the experiment, 100 four-letter words were presented at a 4-sec rate, with the subject instructed to remember them for a later test. Half the words appeared in uppercase letters and occupied $3/8 \times 2 \frac{3}{4}$ in of the display (Capital mode). The other half were in smaller, lowercase letters and occupied $3/16 \times 1 \frac{3}{8}$ in area of the display (Small mode). Half the words, randomly determined, appeared in one mode and half in the other. Following the study phase, 200 words were presented--the 100 old words plus 100 new words, randomly ordered, and subjects were to indicate for each test word whether it was old or new by pushing a button. Reaction times were recorded. Half the subjects were tested with all test words in the Capital mode and half were tested in the Small mode. The interval between the end of the study phase and beginning of the test phase was 3 min.

For each subject, the median correct reaction time was determined for each condition. Means of median times and error percentages are presented in Table 1. The data of primary interest here concern the interaction between presentation and test modes in determining reaction time. The interaction was statistically significant, and is in the expected direction--when a word was presented for test in the same mode as it occurred in originally, recognition was faster than when the two

modes were different ($p < .025$). The interaction in recognition accuracy was in the same direction as that in recognition latency, but was not significant statistically.

Table 1

Mean of Median Correct Reaction Times and
Error Percentages (in Parentheses)

Test Mode (Group)	Presentation Mode			Mean
	Capital	Small	New	
Capital	909 (22.7)	934 (27.0)	1063 (21.2)	992 (23.0)
Small	855 (27.0)	828 (27.9)	939 (17.1)	890 (22.3)
Mean	882 (24.8)	881 (27.4)	1001 (19.2)	941 (22.7)

The experiment shows that the effectiveness with which a recognition test word retrieves a trace of a past occurrence of the word depends in part upon the physical similarity of the test and original stimuli. Visual dimensions of the test word would be expected to play such a role in the recognition decision only if a visual code of the original stimulus were still present at the time of the test. Since the retention interval here was on the average approximately 10.5 min, it appears that perceptual information is longer-lasting than most memory theorists have assumed.

Experiment V: Modality tags and memory for repetitions.

This study was done in conjunction with Richard A. Block and Jeffery J. Summers, and is Experiment I in a paper appearing in the Journal of Verbal Learning and Verbal Behavior, 1973, in press.

When a word list is presented in which some words occur in the visual (V) and others in the auditory (A) modality, subjects can judge the input modality of an individual word fairly accurately even after a delay of several minutes (Bray & Batchelder, 1972; Hintzman, Block, & Inskeep, 1972). Experiments V and VI investigated incidental memory for both input modalities of a word that occurred twice. Both experiments used the same general technique: A long series of words was presented with the subjects instructed simply to remember the words. Some words occurred only once, in either the V or the A modality, while others occurred twice, with the modality of P_2 either the same as or different from that of P_1 . Following presentation, subjects were given a test in which they were to assign each word to one of seven categories --N (not presented), A, V, AA, AV, VA, or VV--according to its history of presentation in the list. From these judgments one can assess memory for modality, for frequency, and in the case of AV and VA, for the order of modalities.

Experiment V used this technique to determine whether or not subjects can remember both modalities of a repeated word. The experiment was originally viewed as a further test of the hypothesis that repetition of a stimulus enhances its retention by producing multiple memory traces--one for each time the stimulus occurred. Experimental support for this notion has been presented by Hintzman and Block (1971). Their experiments showed that traces of different occurrences of a word can be discriminated on the basis of their time tags (the nature of the time tag was investigated in Experiment I). Traces of different occurrences might also be discriminated on the basis of other kinds of information such as input modality. One purpose of this experiment was to determine whether this is the case. Given that subjects do remember both input modalities of a repeated word, two further questions are relevant. First: is the apparent frequency of a word greater if the word occurred in two different modalities than if it occurred twice in the same modality? And second: do subjects retain information about the temporal order of modalities of AV and VA words?

Experimental items were 45 three-letter nouns. Five were randomly assigned to each of nine experimental conditions, representing all combinations of A, V, and N (not presented) in the first and in the second halves of the list. Each condition is identified by a letter pair, with the first letter indicating presentation in the first half and the second in the second half of the list. Thus, words in Conditions AA, AV, VA, and VV occurred twice, once in each half of the list; those in

AN, NA, VN, and NV occurred once; and those in condition NN occurred only on the test form, and not in the list. The list was 80 words long, including 20 filler items evenly divided between the beginning and end of the list. The spacings of repeated words ranged from 24 to 34, and averaged 29 intervening items. A and V presentations were synchronized by means of a tape recorder, Kodak Synchronizer, and Carousel projector. Presentation was at a 3-sec rate. The subjects were 47 paid volunteers, run in groups of up to six subjects each. On a test form on which all 45 words were listed, subjects first judged the frequency of occurrence of each word and then assigned it to one of the seven categories.

Table 2
Response Proportions and Mean Judgments of Frequency,
Experiment V

Condition	Response							Mean Frequency Judgment
	AA	AV	VA	VV	A	V	N	
AA	.39	.12	.13	.02	.23	.05	.06	1.60
AV	.04	.35	.15	.06	.15	.20	.05	1.55
VA	.05	.09	.39	.05	.23	.14	.05	1.53
VV	.01	.13	.11	.33	.05	.28	.09	1.49
AN	.03	.05	.06	.01	.55	.15	.15	1.00
NA	.06	.06	.08	.01	.46	.15	.18	1.03
VN	.01	.04	.04	.02	.15	.52	.21	0.90
NV	.00	.08	.04	.02	.14	.54	.18	0.95
NN	.02	.01	.01	.01	.15	.11	.69	0.36

Table 2 presents the proportions with which the various modality-judgment combinations were given in each of the nine experimental conditions, and also the mean frequency judgment for each condition. First, consider the frequency judgment data (these data can be derived from the modality judgments, as well as from the numerical frequency

judgments). Judgments increased with frequency ($p < .001$) and were essentially unaffected by modality conditions. This is especially interesting for the double-presentation conditions; neither the modality of P_1 nor of P_2 had an effect, and in addition there was no effect of switching modalities vs. keeping them the same ($\chi^2 > .05$ in each case). Apparently the physical similarity of two occurrences of the same word has little effect on the ability of a subject to distinguish between their traces at the time of the test.

The modality judgments of Table 2 indicate that two modality 'tags' can be remembered. In every condition the correct response was most frequent and errors were related to the correct response in a very orderly way. In conditions AV and VA it can be seen that the most likely error, given that the subject said the word occurred twice, was one in which both modalities were given incorrectly (VA for AV and vice versa). This is the kind of error that would occur often if modality tags and time tags were independently associated with the word's meaning. However, it is clear that the order of modalities was usually retained, since for both AV and VA conditions, the correct ordering was given more often than the incorrect ordering.

Experiment VI: Modality tags and repetitions: Locus of the spacing effect.

This study was done in conjunction with Richard A. Block and Jeffery J. Summers. It was the subject of a talk presented at the meeting of the Psychonomic Society, in St. Louis, Missouri, on November 2, 3, and 4, 1972. It is Experiment II in a paper appearing in the Journal of Verbal Learning and Verbal Behavior, 1973, in press.

When P_2 of a to-be-remembered item is given a short time after P_1 , later retention of the item is poorer than it is when the repetition is delayed. This effect of the spacing of repetitions has been found in a wide variety of different memory tasks, using a variety of different experimental materials and several different retention measures (Hintzman, 1974). One fundamental question regarding the spacing effect, which has implications for the various explanations of the effect, is: Which of two presentations, P_1 or P_2 , is the locus of the effect?

According to Hypothesis P_1 , the first presentation is the locus of the effect. This is, retention of P_1 varies with spacing, while retention of P_2 does not. Hypothesis P_2 , by contrast, states that the retention of P_1 is unaffected by spacing, and the locus of the effect is entirely on the retention of P_2 . The problem of differentiating between these two hypotheses can be attacked fairly directly using modality judgments. If traces of two occurrences are differentially tagged with modality information, it should be possible to tell from modality judgments which of the two is better retained as spacing increases. This was the primary purpose of Experiment VI.

The materials and procedure were basically the same as in Experiment V. A total of 96 experimental words were used. Within the list, 24 of the words did not occur (Condition N), and 12 each were assigned to Condition A, V, AA, AV, VA, and VV. Within each of the latter four conditions, three words were repeated at each of four spacings, $S = 0, 1, 5$, and 15 intervening items. On the test sheet 72 words appeared--all 48 double-presentation words plus half the words in each of Conditions A, V, and N. The presentation list was 140 items long including fillers. The subjects were 191 paid volunteers.

In Figure 3 are presented the mean frequency judgments as a function of modality conditions and spacing. Analysis of variance on these data indicated that the spacing effect was significant for each condition individually, and there were no significant interactions with conditions (all interaction F 's < 1). The most important of these tests indicates that the size of the spacing effect is the same regardless of whether the modality of P_2 is the same as that of P_1 or different. This indicates that the spacing effect is produced at a semantic level of analysis.

In order to determine the locus of the spacing effect, frequency judgments for conditions AV and VA were broken down into two components, one for A judgments and one for V judgments. (For a complete explanation

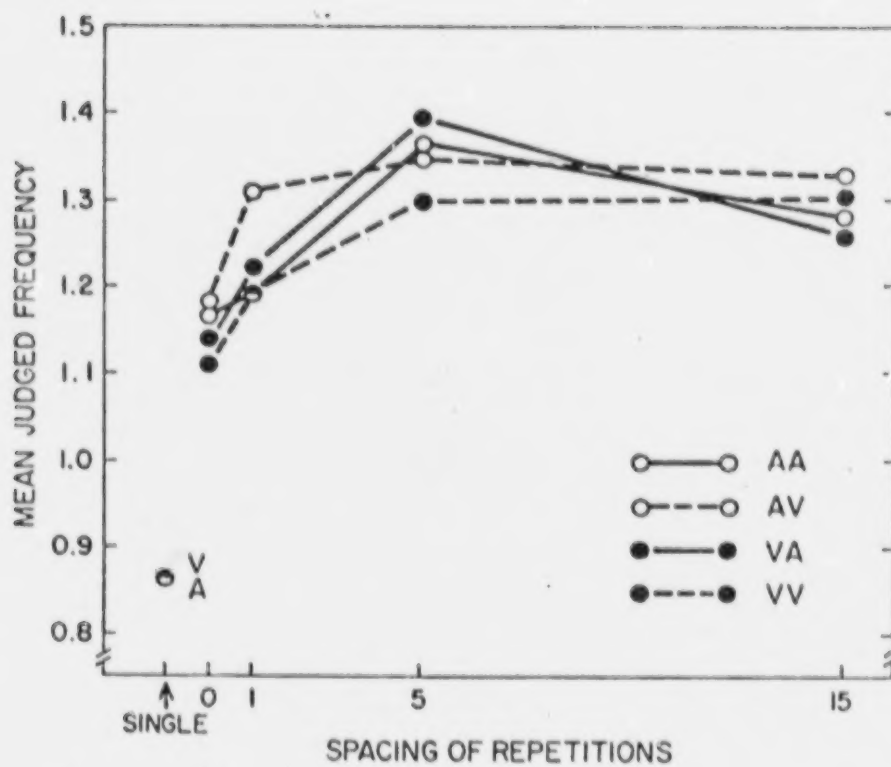


Figure 3. Mean judgments of frequency as a function of spacing and modality condition; Experiment VI.

of this analysis, see Hintzman, Block, & Summers, 1973). The sum of the two components for a given condition equals the mean judgment of frequency. In Figure 4, the A and V components are plotted as separate functions of spacing for Conditions AV (top panel) and VA (bottom panel). In either graph the A component is represented by open and the V component by closed circles. It can be seen that in both conditions the magnitude of the spacing effect was greater on judgments corresponding to the modality of P_2 than on those corresponding to the modality of P_1 .

It may appear that these data are inconsistent with either Hypothesis P_1 or Hypothesis P_2 , since an effect of spacing was obtained on both components of judged frequency. However, neither hypothesis predicts that one component will reflect the entire spacing effect, since modality is not perfectly remembered. An increase in retention of a V presentation should lead to an increase in both V and A judgments, but the increase in the latter should be smaller, since it is due to generalization. The degree of generalization between A and V components can be estimated from the single-presentation data (again, see Hintzman, Block, and Summers, 1973, for details), and quantitative fits of Hypotheses P_1 and P_2 can be made to the obtained component data. When this was done, it was found that the mean absolute prediction error for Hypothesis P_2 was .019 for Condition AV, and .003 for Condition VA. This fit was better than that of Hypothesis P_1 (.035 and .024, respectively) and also than that of the hypothesis that the spacing effect falls equally on P_1 and P_2 (.026 and .013, respectively). Thus the hypothesis that it is retention of P_2 that increases with spacing, while that of P_1 is unaffected, fits the data from both conditions quite well. The solid curves in Figure 4 are the predictions of Hypothesis P_2 .

A final point about the outcome of Experiment VI is that the ability of subjects to remember the order of modalities in Conditions AV and VA did not vary with spacing. This indicates that the order of modalities is probably not judged on the basis of time tags, since if it were performance would deteriorate as P_1 - P_2 spacing became shorter. Instead, it may be that order information is encoded in the same way as spacing information is encoded (see Experiment II). When P_2 occurs, it retrieves the trace of P_1 , so that the subject is simultaneously aware of both modalities and of their order. If this information is stored, then on a later judgment test it can serve as the basis of a judgment of the order in which the A and V presentations occurred.

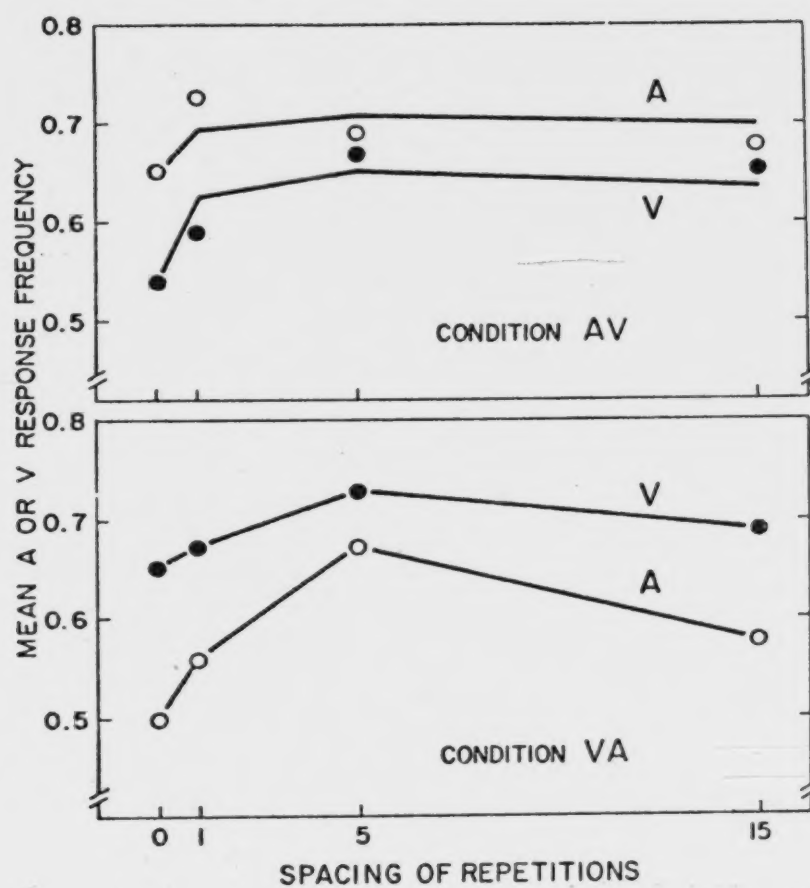


Figure 4. Component analysis of judgments for Conditions AV and VA; Experiment VI. Solid curves are predictions of Hypothesis P_2 .

Experiment VII: Spacing effects in picture memory.

This study was done in conjunction with Miriam K. Rogers, and is Experiment I in a paper appearing in Memory & Cognition, 1973, in press.

A number of theoretical explanations of the spacing effect have been proposed. The more prominent theories are reviewed in a paper presented by the present investigator at the 1973 Loyola Symposium on Cognition (Hintzman, 1974). In one of the most popular theories of the spacing effect, rehearsal plays a crucial role (Atkinson & Shiffrin, 1968; Greeno, 1967; Waugh, 1970). By rehearsal we mean here the voluntary retrieval and reprocessing of a memory trace when the stimulus it represents is no longer physically present. According to the rehearsal hypothesis, an item for which P_2 occurs shortly after P_1 is rehearsed less, and hence is remembered less well on a later test, than is an item for which the P_1 - P_2 interval was long. Direct support for the rehearsal hypothesis comes from the experiments of Rundus (1971), who studied the rehearsal patterns of subjects who were asked to rehearse aloud during presentation of a free recall list. When the spacing of repetitions was varied, less overt rehearsal was observed of words that had short spacing intervals than of words that had long spacing intervals. Rundus's data indicate, in addition, that the differential rehearsal took place entirely during the P_1 - P_2 interval; suggesting that the spacing effect may have a rather trivial explanation. Perhaps long spacing intervals lead to better retention simply because they provide more time to rehearse P_1 before P_2 occurs.

This account seems to contradict Experiment VI, which appears to show that it is retention of P_2 , not of P_1 , that increases with spacing. However, one could suppose that the subject continues rehearsing P_1 even after P_2 occurs when the spacing interval is short. Thus the previous experiment does not necessarily contradict the rehearsal hypothesis.

The purpose of this experiment was to test the rehearsal hypothesis in another way--by using stimulus materials that subjects do not rehearse. Shaffer and Shiffrin (1972) have reported that recognition memory for complex visual scenes is affected by stimulus "on" time, but not by the blank "off" time following presentation. From this they concluded that subjects do not rehearse such pictures. The stimulus materials chosen for this study, therefore, were colored slides. In order to avoid ceiling effects in recognition memory for pictures, judgments of frequency were used as the dependent variable. If the spacing effect is caused by rehearsal, then judged frequency of pictures should not show a spacing effect. This prediction was tested in Experiment VII.

The experimental stimuli were 120 color vacation slides, selected to minimize inter-item similarity. Thirty scenes were assigned at random to each of four frequencies of occurrence ($F = 0, 1, 2$, and 3); and within the $F = 2$ and 3 conditions, 10 scenes were assigned at random to each of three spacings ($S = 0, 1$, and 5 intervening items). For $F = 3$ items, the

P_1 - P_2 and P_2 - P_3 spacings were the same. The list consisted of 190 slides, including 10 filler slides. The subjects were 27 paid volunteers, run in three groups of up to 10 subjects each. The list was presented at a 3-sec rate with subjects instructed to remember the pictures for a later test. Then the pictures were presented one at a time at a 5-sec rate, and subjects wrote their frequency judgments on a test sheet.

The mean judgments of frequency are presented in Figure 5. Judgments increased as a function both of frequency and of spacing, and the magnitude of the spacing effect was greater the more times the pictures occurred. Statistical analyses bore out these conclusions (both p 's $< .001$). The broken line indicates perfect performance, and comparison with the data points shows that frequency judgments were quite accurate. Accuracy was poor only when a picture was repeated at a short spacing ($S = 0$ or 1). This pattern of results is essentially the same as that obtained using verbal materials. Thus, even if one is unwilling to accept the strong conclusion of Shaffer and Shiffrin (1972) that pictures cannot be rehearsed, and admits only that their rehearsal is inefficient compared to words, the magnitude of the obtained effect appears to rule out an explanation in terms of rehearsal.

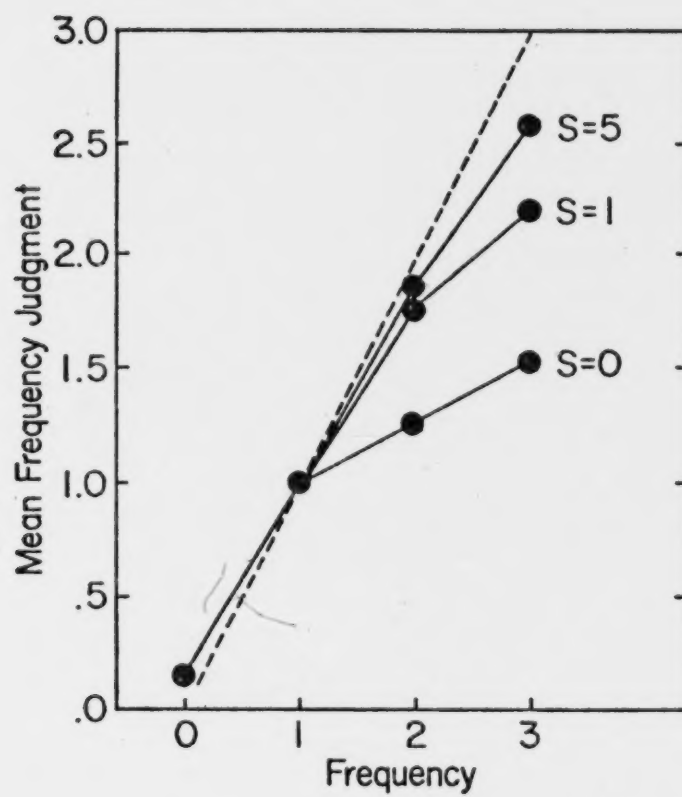


Figure 5. Mean judged frequency of pictures as a function of frequency and the spacing of repetitions; Experiment VII.

Experiments VIII and IX: Time vs. items as determinants of the spacing effect.

These experiments were done in conjunction with Miriam K. Rogers and are experiments II and III in an article appearing in Memory & Cognition, 1973, in press.

The conclusion of Shaffer and Shiffrin (1972) that pictures are not rehearsed also led to their use in Experiments VIII and IX. The purpose of these studies was to determine whether it is the other stimuli that are presented during the P_1 - P_2 interval or the duration of the interval that causes the spacing effect. This problem is a difficult one to investigate using verbal materials, since any empty time can be used by the subject for rehearsal, either of the immediately preceding stimulus or of earlier stimuli. Not only are the temporal conditions of practice outside the experimenter's direct control, but the independent manipulation of P_1 - P_2 time vs. P_1 - P_2 items results in different degrees of learning, making the forms of the spacing curves difficult to compare (e.g., Melton & Shulman data, reported in Melton, 1970). When pictures are used, eliminating rehearsal, presentation rate and degree of learning can be manipulated independently.

Experiments VIII and IX used the same materials and essentially the same method as Experiment VII. They differed from Experiment VII only in using blank (opaque) slides at certain points in the presentation list so that the inter-stimulus interval could be manipulated independently of stimulus duration. The basic idea was to compare effects of P_1 - P_2 intervals filled with presentation of other pictures with effects of intervals of the same length filled with blank slides (no pictures). The differences between Experiments VIII and IX are slight and need not be described here (for details see Hintzman & Rogers, 1973). It should be sufficient for purposes of this report to simply state the conclusion: while in both experiments there was a slight effect on judged frequency of a picture of other pictures occurring during the P_1 - P_2 interval, time itself was far more important as a determinant of the spacing effect. In fact, the data from all three studies (VII, VIII and IX) were similar enough to be combined, as they have been in Figure 6. In this figure, frequency judgments from all three experiments are plotted as a function of the duration of the spacing interval. Filled intervals are represented by closed circles, and unfilled intervals (blank intervals) are represented by open circles. The upper curve is for pictures which occurred three times, and the lower one for pictures occurring twice. It is clear that the effect of spacing on judged frequency was primarily a function of time--particularly for the pictures that occurred two times. Thus a major conclusion of these studies is that the spacing effect--at least where pictures are concerned--is primarily determined by time, and not by the information-processing activities of the subject during the spacing interval. A secondary conclusion is that since unfilled intervals do not lead to higher judgments of frequency than filled intervals, there was no effect on judged frequency that could be attributed to facilitatory rehearsal. This conclusion is in agreement with that of Shaffer and Shiffrin (1972) that pictures are not rehearsed. It thus reinforces the interpretation previously given for Experiment VII.

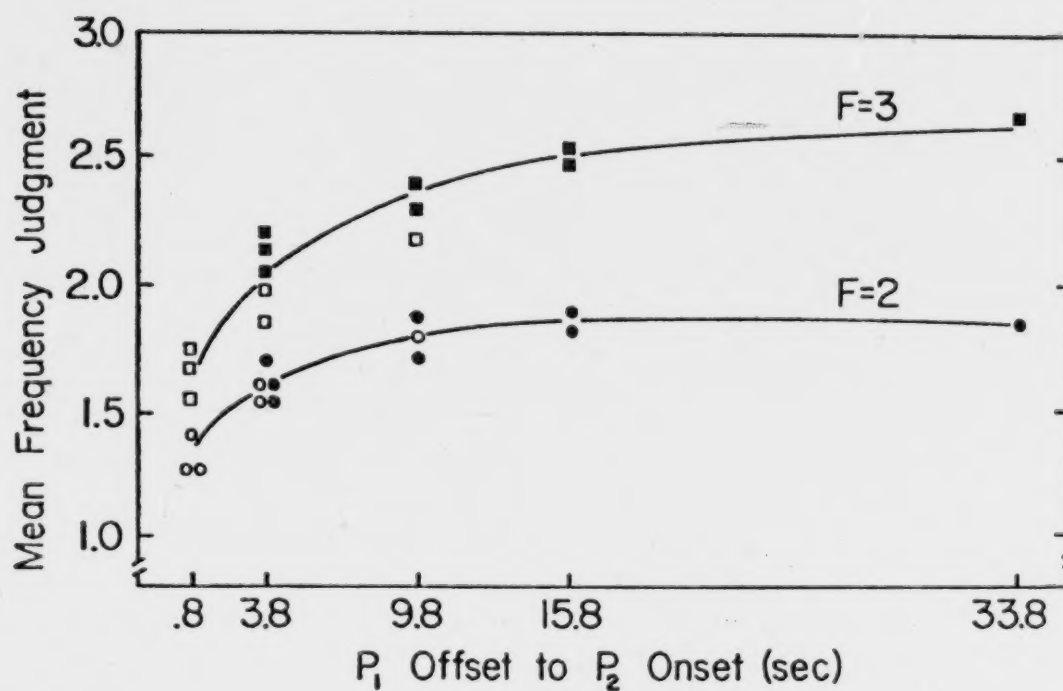


Figure 6. Combined data of Experiments VII, VIII, and IX: Mean judged frequency as a function of duration of the spacing interval.

Experiments X and XI: Attention and the spacing effect: Physiological measures.

A major class of theories of the spacing effect reviewed by Hintzman (1974) attribute the effect to attention mechanisms. It is assumed that less attention is paid P_2 when it occurs shortly after P_1 than when the P_1 - P_2 interval is longer. Attention explanations of the spacing effect have the advantage that they unequivocally predict the conclusion of Experiment VI, that it is the retention of P_2 , rather than that of P_1 , that suffers from repetition at short lags. They also have the advantage of being able to explain the outcome of Experiment VII, which showed that a spacing effect is found when pictures are the to-be-remembered items. In discussing Experiments XII, XIII, and XIV, we will make a distinction between voluntary attention mechanisms and involuntary attention mechanisms. For the purpose of discussing Experiments X and XI, however, the distinction is not necessary.

Both these experiments were aimed at providing evidence for an attention explanation of the spacing effect through measurement of physiological correlates of attention. Experiment X was done in conjunction with Michael I. Posner, and its purpose was to determine whether evoked EEG potentials elicited by P_2 of a word were higher when the P_1 - P_2 interval was long than when it was short. There is evidence of a decrease in evoked potential when a stimulus is repeated immediately in certain situations, but no evidence that this would be found with repetition of a word under conditions typical of memory experiments. EEG recordings were taken while subjects studied a 260-item word list presented at a 3-sec rate. There were two experimental conditions, intermixed in the list. In one, words occurred four times in succession. In the other, they occurred four times at a spacing of four intervening items. After several subjects had been run, evoked potentials were averaged and a comparison was made between P_1 and P_4 potentials for both the Massed and Distributed conditions. No effects of massed repetitions were evident, and so the experiment was abandoned.

Experiment XI involved a similar effort to measure attention, using pupil dilation as the dependent variable. During his year spent at the Oregon Research Institute, Daniel Kahneman set up his pupillometry apparatus, which we were invited to use. The apparatus can be used to photograph the eye of the subject at precisely-determined points during the experiment. Previous experiments have shown that increases in pupil size occur in situations in which voluntary mental effort is high, suggesting that pupil dilation may be a correlate of attention (Goldwater, 1972). In Experiment XI, subjects were presented auditorily with several lists of words which they were asked to free recall. Words were presented at a 4-sec rate, and photographs were taken every 1 sec. Repeated words had spacings of 0, 1, and 4 intervening items. We hoped to find that the degree of pupil dilation produced by P_2 increased with P_1 - P_2 spacing. After five subjects had been run, a preliminary analysis was done. Although the data indicated a slight tendency for the pupil to dilate when a word was presented, it was much too small and the data were much too variable to allow any differences in dilation of P_2 as a function of

spacing to be observed. We concluded that the free recall task probably does not impose the mental load necessary to strongly affect pupil size. Because the obtained dilation was small and because of the fact that pupil measurements with our apparatus are extremely time consuming to make, the experiment, like Experiment X, was abandoned as unpromising.

Experiment XII: The spacing effect and control of attention through incentive.

The failures of Experiments X and XI to demonstrate the role of attention in the spacing effect by taking physiological measures of attention do not necessarily indicate that some other explanation of the effect is required. Experiment XII attempted to demonstrate the role of attention in a more direct way--by controlling the degree of attention the subject gave the to-be-remembered item.

For present purposes, it is useful to distinguish between voluntary and involuntary attention mechanisms. This experiment had to do with voluntary attention. By this is meant the subject's voluntary direction of processing effort to a stimulus event. An involuntary attention explanation of the spacing effect is the subject of Experiments XIII and XIV.

How can a voluntary process be distinguished from an involuntary one? Both must be lawfully related to various independent variables--the difference has to do with the kinds of independent variables that affect each. Voluntary processes are highly flexible, and are governed by the subject's beliefs, expectations, and past experience in the task. Involuntary processes, on the other hand, are relatively inflexible and can be changed by experience only gradually if at all. As suggested by Hintzman (1974), a fairly straightforward operational distinction between voluntary and involuntary processes can be formulated in terms of instructional variables. Assuming that a subject's beliefs and expectations can be controlled through experimental instructions, an experimental outcome that has been demonstrated to be vulnerable to instructional manipulation may be said to be voluntary. One that cannot be so manipulated even by instructions specifically aimed at altering the voluntary process imagined to be responsible, may be said to be involuntary. Convincing evidence for a voluntary attention explanation of the spacing effect, according to this analysis, must rest on demonstrations that the effect can be eliminated or sharply attenuated by appropriate experimental instructions or other manipulations of the subject's strategies.

This reasoning was behind Experiment XII, done in conjunction with Jeffery J. Summers. The purpose was to test the voluntary attention hypothesis by manipulating the attention given to second occurrences of repeated items. The to-be-remembered materials were 193 color vacation slides. The list was presented by a projector synchronized with a tape recorder in such a way that, where appropriate, a tone could be presented simultaneously with the onset of a picture. Fifty pictures were assigned to Condition N (not presented); 40 occurred one time only, and half of these were accompanied at onset by an audio tone and half were not. Eighty additional pictures were presented twice each, 20 at each of four levels of spacing ($S = 0, 1, 5, \text{ and } 15$). Ten of the pictures at each spacing were accompanied at P_2 onset by a tone, and 10 were not. None of these repeated pictures was accompanied by a tone at P_1 . The subjects were 45 paid volunteers. They were told to remember the slides they would be shown for a later test, and that they would be paid for high accuracy. They were further instructed that pictures accompanied by a tone would

be worth 4¢ on the test, while those not accompanied by a tone would be worth only 1¢. Thus incentive value was manipulated in a way designed to produce differential attention to P_2 in the Tone and No Tone conditions. If a flattened spacing curve resulted from presentation of the tone at P_2 (high attention), the outcome would support a voluntary attention hypothesis. Pictures were presented at a 3-sec rate. On the final test, 170 pictures were shown, and subjects gave judgments of frequency. Subjects were paid either 4¢ or 1¢, as had been indicated in the earlier instructions, for each picture--but payment was contingent on a correct judgment of frequency.

The results are shown in Figure 7. The two data points for single presentation pictures indicate that those occurring once accompanied by a tone were given higher frequency judgments than were those occurring without the tone. Thus, the incentive manipulation apparently did affect attention. The two spacing curves show that a tone at P_2 had a similar effect, so the manipulation of attention given to P_2 was apparently successful as well. However, the effect of the tone did not interact with the effect of spacing ($F < 1$). If spaced repetitions are ordinarily given nearly maximum attention and massed repetitions are given very little, then a manipulation of attention paid P_2 would most likely have its greatest effect at shorter spacing intervals. The failure to find an attenuation of the spacing effect when P_2 was accompanied by the tone is not conclusive; but an obvious interpretation is that the spacing effect involves a different mechanism than the one that was affected by our instructions regarding differential payoffs.

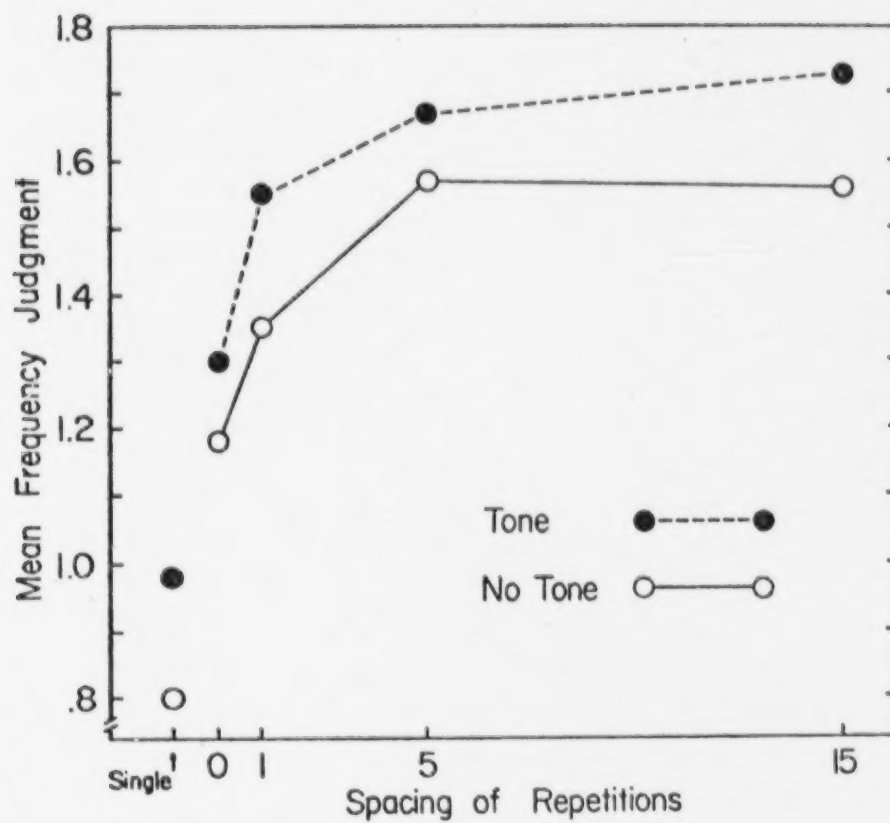


Figure 7. Effects of spacing and the manipulation of attention given P_2 on judged frequency of pictures; Experiment XII.

Experiments XIII and XIV: On habituation and recovery as an explanation of the spacing effect.

The habituation explanation of the spacing effect is like the voluntary attention hypothesis in that it explains the effect as due to deficient registration of P_2 when the P_1 - P_2 interval is short. The idea is that when the to-be-remembered stimulus occurs, some internal representative of the stimulus, the activation of which is necessary for an enduring memory trace to be stored, becomes adapted--that is, its threshold for activation is raised (Hintzman, 1974). This could be interpreted in an all-or-none learning framework as a decrease in the probability of storage of a new trace, or in an incremental framework as a decrease in the strength of any new trace that is stored. Adaptation or habituation may be assumed to continue for as long as the stimulus is present, and recovery to begin when the stimulus ceases. The spacing effect follows directly from these assumptions; if P_2 comes before recovery from the effect of P_1 is complete, later retention will be poorer than it will be if P_2 is delayed. The analogy with the behavioral habituation studied in lower animals, suggested by the present choice of terms, should be considered only a rough one. In the present use of the term, it is the 'response' of storing a particular kind of memory trace that habituates.

Experiment XIII tested one prediction of the habituation-recovery hypothesis. If habituation is assumed to continue the longer the stimulus is studied, then recovery should presumably take longer, the longer was the exposure duration of P_1 . In this experiment, three durations of P_1 were used: 2.25, 5.25, and 8.25 sec onset-to-offset times; and these were combined orthogonally with 6 different spacing intervals, of 0.75, 3.75, 6.75, 9.75, 12.75, and 18.75 sec from P_1 offset to P_2 onset. The duration of P_2 was always 5.25 sec. In addition, there were three conditions in which each item occurred only one time, for durations of 2.25, 5.25, or 8.25 sec. Stimuli were color vacation slides, and six were assigned to each of 22 conditions--the 21 defined above and an $F=0$ condition. The entire presentation list consisted of 290 slides, including 20 fillers inserted at the beginning and end of the list.

The subjects were 60 paid volunteers. As in previous experiments, they were told to study the pictures and remember them for a later test. Picture durations were controlled by means of a Kodak Synchronizer triggered by pre-recorded signals on an audio tape. On the final test, subjects gave frequency judgments, and the test stimuli were presented at a 9-sec rate.

The data are shown in Figure 8. It can be seen that for pictures occurring one time only, exposure duration affected judged frequency to some extent ($p < .01$); and duration of P_1 increased the judged frequency of pictures occurring twice ($p < .005$). However, the effect of P_1 duration did not interact with that of spacing ($F < 1$). Thus there was no evidence for the expected effect of P_1 duration on recovery time. Spacing appears to have been determined only by P_1 offset to P_2 onset time.

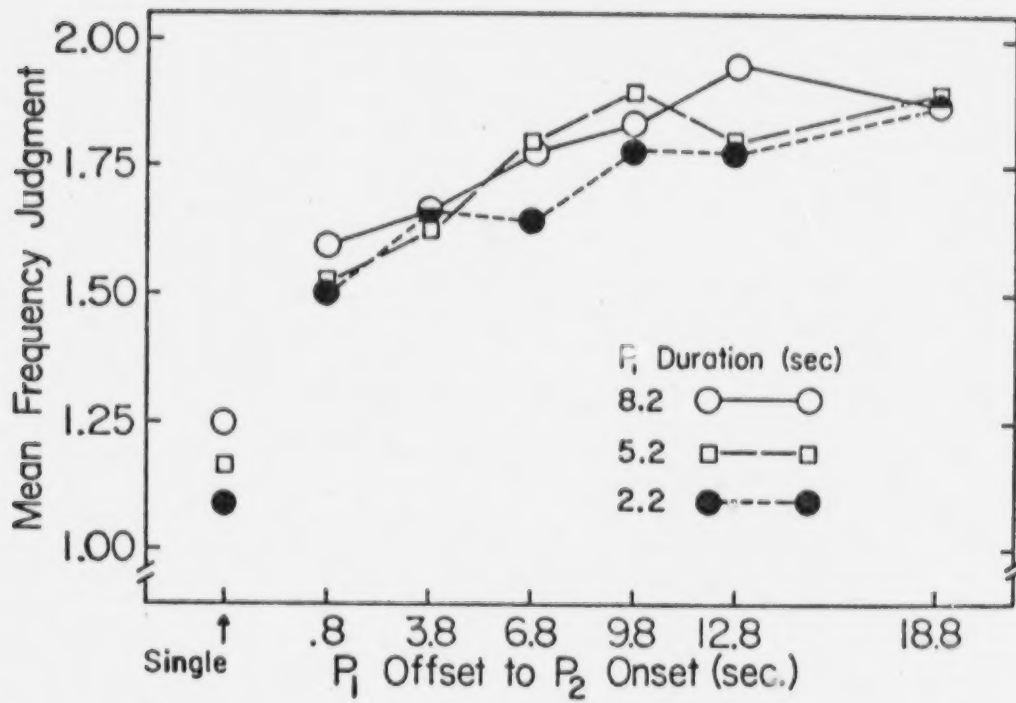


Figure 8. Effects of spacing and P_1 duration on judged frequency of pictures; Experiment XIII.

It seems possible that habituation asymptotes after about 2 sec of study of a single stimulus, and that it is repetition, rather than duration, that produces effective habituation. For this reason, in Experiment XIV the manipulation of P_1 duration in Experiment XIII was replaced by 1, 2, or 3 prior massed presentations. There were 14 conditions: one in which the pictures did not occur, one in which they occurred one time each, and 12 in which they occurred more than once. The multiple-presentation conditions involved all combinations of 1, 2, or 3 prior massed presentations (M) and 0, 1, 3, and 7 intervening items in the final ($P_M - P_{M+1}$) spacing interval. Thus, in the $M = 3$ conditions for example, P_1 , P_2 , and P_3 occurred in close succession (massed repetition) and the $P_3 - P_4$ interval varied in length. There were 28 pictures assigned to each of the 14 conditions, plus some filler items which occurred several times and did not fit this pattern. The subjects were 41 paid volunteers. Presentation was at a 3-sec rate.

The results are shown in Figure 9. Frequency judgments increased as a function of M , as would be expected since $M = 1, 2$, and 3 conditions correspond to $F = 2, 3$, and 4, respectively. However in all conditions, all presentations except the final one (P_{M+1}) were massed. The important result is that the form of the spacing curve did not interact with M ($F < 1$). Experiment XIV thus confirmed the outcome of Experiment XIII. The spacing effect seems to be primarily a function of P_M offset to P_{M+1} onset time. Neither the duration of a single prior presentation nor the number of prior presentations affects the form of the spacing curve. Thus, if the habituation-recovery hypothesis is the correct explanation of the spacing effect, it is difficult to obtain direct evidence for the process. Unlike behavioral habituation, the recovery rate here is independent of the amount of experience with the stimulus.

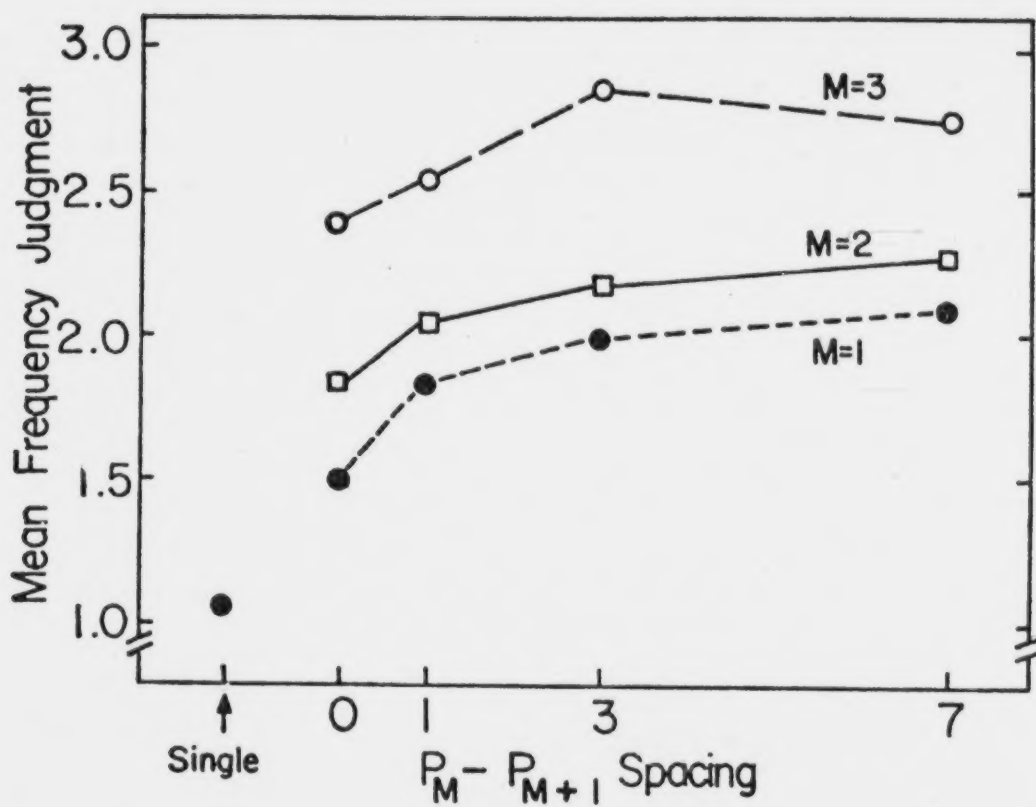


Figure 9. Effects of the number of prior massed presentations (M) and spacing of the last presentation on judged frequency; Experiment XIV.

Experiment XV: Effect of spacing on storage of P_2 : All-or-none or incremental?

In Experiment VI, evidence was presented suggesting that it is retention of P_2 that suffers when a to-be-remembered item is repeated at a short spacing. There are at least two possible interpretations of this conclusion regarding the representation of P_2 in memory. One is that encoding of P_2 is all-or-none, and at short spacings the probability of encoding P_2 is lower than it is at long spacings. The other is that at short spacings the trace that is stored representing P_2 is weak, and that the longer the P_1 - P_2 interval becomes, the stronger is the trace of P_2 that can be stored. While most theories of the spacing effect seem to adopt the latter, incremental, point of view, there is no good evidence favoring it over the all-or-none interpretation. And at least one memory model which is easily adapted to predictions in the frequency-judgment task assumes that different presentations can result in different memory traces, but the existence of a trace of a given presentation is an all-or-none matter (Bernbach, 1970).

Experiment XV attempted to determine which view, all-or-none or incremental, is correct. Pictures were presented and spacing varied, as in previous experiments, and on the test subjects were asked to give, in addition to a frequency judgment for each picture, a numerical confidence rating indicating how certain they were that the frequency judgment was correct. The primary question concerned confidence ratings given to $F = 2$ items when the P_1 - P_2 interval is short. It was reasoned that if the all-or-none hypothesis is correct, then the subject's confidence in his frequency judgment should not be affected by spacing. When the subject correctly judges the frequency as two, he should be just as certain at short as at long spacings. And likewise, when he incorrectly judges the frequency to have been one, he should be as confident as when the frequency actually was one.

The experimental items were color vacation slides. There were five conditions: in one were 45 pictures that were not presented in the list; in another there were 45 pictures which occurred once; and 20 pictures, in each of three other conditions, occurred twice, with spacings of $S = 0, 1$, and 5 intervening items. The entire presentation list, including filler items was 175 items long. The subjects were 26 paid volunteers, and the presentation rate was one slide every 3 sec. On the test, subjects were given 9 sec per item to respond. For each picture they first circled a frequency judgment (0, 1, or 2), and followed it with a confidence rating for the frequency judgment. The confidence scale was labeled, right to left, as follows: Very Uncertain; Somewhat Uncertain; Pretty Certain; Absolutely Certain. Subjects expressed confidence by placing an X in one of the four categories.

Mean frequency judgments increased with both frequency and spacing, as in previous experiments. Of interest here are the mean confidence ratings for different judgments of frequency in the different conditions.

These data, together with the percentage of times each frequency judgment was given, are presented in Table 3. Two things to note about these data concern the effects of frequency and of spacing. First, reading down columns of the table, it can be seen that on the average subjects were more confident in a particular frequency judgment when it was correct than when it was incorrect. This was true for all three judged frequencies, and indicates that subjects actually are able to monitor the accuracy of their frequency judgments. These confidence ratings cannot be based directly on strength or number of traces, since the co-variance with judged frequency is far from perfect. Second, note the effect of spacing. As spacing between P_1 and P_2 increased, confidence in judgments of zero decreased, as did those of judgments of one. However, subjects became more confident that a frequency judgment was correct when it was a judgment of two.

Table 3

Mean Confidence Ratings and Response Percentages, Experiment XV.

Frequency Spacing		Judged Frequency					
		0	(%)	1	(%)	2	(%)
0	-	3.21	83	2.23	15	3.00	1
1	-	2.95	28	3.16	60	3.19	12
2	0	2.85	17	3.05	49	3.38	35
	1	2.82	14	2.94	38	3.70	48
	5	2.35	13	2.86	30	3.70	57

This is exactly the outcome that was thought to be contrary to the all-or-none view, since it suggests that there is partial information about P_2 when the spacing was short. Further thought, however, has convinced us that while there are some all-or-none models that are inconsistent with this outcome, there are others with which it is compatible. Like other attempts to decide between incremental and all-or-none views, this one appears to have been inconclusive.

Experiments XVI and XVII: Memory and the experience of duration in retrospect.

These two studies appear as Experiments I and II in a doctoral dissertation done by Richard A. Block under the direction of the principal investigator. A revised version of the dissertation is to be published in Memory & Cognition (in press).

How does one remember the length of an interval of time? A number of hypotheses have been offered--all involving memory. The most specific hypothesis has been offered by Ornstein (1969). It states that the experience of duration in retrospect is a cognitive construction based on the 'storage size' of the interval, as assessed by the subject from memory at the time the judgment is made. Storage size depends on (a) the number of events stored and retrieved, and (b) the complexity of the coding of the events at the time of retrieval. Experiments XVI and XVII attempted to test certain aspects of Ornstein's notion, and to determine whether some common measure of memory might be related to judgments of duration in a fairly direct way.

Both experiments used essentially the same method. The subjects attended to a sequence of events consisting of three segments: a standard interval of music, an experimental interval during which a list of words was presented, and a second standard interval of music. In both experiments there were two conditions, differing only in the events which occurred during the experimental interval (the word list). After attending to the three intervals, the subjects were asked to make three comparative judgments of duration, with each of the three intervals being compared with each other interval. Finally, the subjects were given three memory tasks designed to measure memory for the events that occurred during the experimental interval: free recall, recognition, and judgment of the number of events that occurred.

Experience of duration in retrospect is usually found to increase as a function of the number of events that occurred during the interval. Experiment XVI was designed to investigate this phenomenon in a situation in which two conditions expected to lead to the same degree of retention (as assessed by free recall) involve the presentation of different numbers of events. If the total time taken for presentation of two word lists is the same but the number of words presented is different, free recall performance is usually about the same for both lists (the 'total time hypothesis'; cf. Cooper & Pantle, 1967). In this experiment, the experimental interval was a 180-sec series of auditorily presented words. In one condition, there were 60 words read at a 3-sec rate. In the other, there were 30 read at a 6-sec rate. The subjects were 66 paid volunteers; half were run in the 60-word condition and half in the 30-word condition. Briefly, the results showed that, in keeping with earlier experiments, the subjects who had been read 60 words remembered the experimental interval as longer than the subjects who had heard 30 words. The free-recall data, however, were consistent with the total time hypothesis--the 60-word subjects recalled

an average of 11.0 words while the 30-word subjects recalled 10.9 ($p > .05$). Subjects in the 60-word condition thought more words had been presented during the interval; and recognized individual words more poorly but could recognize more words altogether than subjects in the 30-word condition. It appears that number of words that can be recalled from an interval is not the primary determinant of memory for duration of the interval.

If memory for duration is not based on event recall, perhaps it is based on complexity. In Experiment XVII, two conditions were compared which differed in the complexity of the word list. If memory for duration is based on complexity, then the high-complexity subjects should remember the experimental interval as longer than do the low-complexity subjects. At the same time, however, the low-complexity condition is known to lead to higher free recall performance than the high-complexity condition. So amount recalled and complexity would not be expected to work in the same direction.

The two conditions were Blocked vs. Random presentation of related words. During the experimental interval, subjects studied 20 slides, presented for 8 sec each. Both groups of subjects were shown the same words, but in the Blocked condition the four words on a given slide were from the same taxonomic category, while in the Random condition the four words were always from four different categories. Presumably, the complexity of the experimental sequence should be higher for the Random subjects than for the Blocked subjects. The subjects were 52 paid volunteers; half served in one condition and half in the other.

The result was contrary to the "complexity" notion: The Blocked presentation subjects judged the experimental interval to be slightly longer than did the Random presentation subjects. The effect was small, but significant statistically ($p < .05$). As expected, free recall was much better for the Blocked condition than for the Random condition: 16.7 and 9.1 mean words recalled, respectively. Judgment of the number of words presented did not differ. And recognition memory was better for the Blocked than for the Random group. Experiment XVII thus seems to rule out complexity as the primary determinant of remembered duration. Apparently some factor or combination of factors other than free recall or complexity alone, determines remembered duration. Beyond this conclusion, these experiments shed no further light on the nature of the process.

CONCLUSIONS

The experiments conducted as part of this grant project were rather far-ranging in terms of the theoretical questions under consideration, but had a great deal in common methodologically. All involved presenting subjects with lists--either of words or of pictures--to memorize. While two (Experiments X and XI) attempted to measure physiological correlates of encoding processes during the study phase of the experiment, the others were concerned directly only with information that could be retrieved on a later test. Of these, one (Experiment IV) used as the dependent variable recognition latency; in the others the primary dependent variables were memory judgments. Judged frequency was the most common measure, but judgments of serial position (Experiment I), of spacing (Experiment II), of mode of input (Experiments III, V, and VI), and of duration (Experiments XVI and XVII) were also used. Conclusions drawn from the 17 experiments reported here can be summarized as follows:

Experiment I was concerned with memory for serial position. Evidence was presented suggesting that the 'time tag' upon which judgments of serial position are based consists of associations between the to-be-remembered event and the cognitive context prevailing at the time it occurred. In addition, however, there is a component of serial-position memory that declines relatively quickly over time, and produces a recency effect much like that found in free recall. This component may be a short-term strength, or may depend on the similarity of the retrieved context and the context prevailing at the time of the test (see Hintzman, Block, & Summers, 1973).

Experiment II showed that subjects are not only able to judge the spacing of repetitions of the same word, as had been shown previously (Hintzman & Block, 1973), but also to judge the spacing of pairs of associatively related words. They are not able to do so for pairs of unrelated words. The finding confirms the hypothesis that judgments of spacings are based on an implicit judgment of recency, stored at the time the second member of the pair is presented, which is retrieved on the later judgment test. The confirmation opens up the possibility of using spacing judgments as indicators of spontaneous retrieval taking place during the study phase of an experiment. Retrieval processes operating during the experimental study phase should therefore receive more attention in future work.

Experiment III extended previous findings regarding the encoding of the A-V distinction and distinctions within the visual mode, by showing that some words are presented by a male and some by a female voice, immediate free recall of the words is clustered by voicing, and voicing is remembered to some extent over a period of several minutes. This still leaves open the question of whether input mode is represented as an abstract proposition, associated with the meaning of the word, or more directly, as a literal copy of the experience.

Experiment IV attempted to answer this question using visual materials and measuring recognition latencies for test words that either matched or mismatched the original stimuli along physical dimensions. It was found that recognition was faster when the original and test stimuli were identical (the meanings were the same in either case), indicating that physical attributes of the original stimulus do indeed endure and play a role in retrieval even after an interval of minutes. The conclusion is contrary to many current theoretical notions which hold that physical attributes of the stimulus are encoded only in sensory memory, which decays within a fraction of a second; or in a modality-specific short-term store which decays within several seconds. Apparently, the transformations undergone by a verbal stimulus (visual to phonemic and semantic codes) do not destroy the visual information--it remains in memory just as the visual memory of a complex pattern does--but simply obscure its existence through the dominant role they usually play in retrieval.

Experiment V showed that modality judgments could be used to investigate the traces of repetitions of the same word. It showed that subjects do remember both modalities of a repeated word, and that they can recall the order when the two modalities were different.

Experiment VI used the conclusion of Experiment V to investigate the spacing effect. In order to determine which presentation, P_1 or P_2 , is the locus of the effect, the two presentations were differentially tagged with modality information, and modality judgments were taken. The results showed that the spacing effect is equally great whether P_1 and P_2 are in the same or different modalities, and that it is apparently later retrieval of P_2 , rather than P_1 , that suffers when spacing is short.

Experiment VII tested the hypothesis that rehearsal processes are the cause of the spacing effect by using scenic pictures, which subjects do not rehearse (Shaffer & Shiffrin, 1972). Judged frequency of pictures was affected by spacing, just as is the case with verbal materials. While one might argue that pictures are in fact rehearsed, so that the experiment is inconclusive, the outcomes of Experiments VIII and IX indicate that they are not rehearsed. The cause of the spacing effect in picture memory could be different, of course, from its cause when verbal materials are used. But there is other evidence against rehearsal as the cause of the spacing effect in verbal memory (Bjork & Allen, 1970); and in any event it can be claimed on the basis of the present work that rehearsal is inadequate as a general explanation of the spacing effect.

Experiments VIII and IX used picture memory to determine whether the spacing effect is primarily a function of the time or the other stimuli occurring during the P_1 - P_2 interval. The results suggest that it is primarily time, not items, that determines the spacing effect. Apparently, when pictures are used the effect is not influenced in any important way by information-processing activities the subject engages in during the spacing interval.

Experiments X and XI sought evidence for the attention explanation of the spacing effect by measuring physiological correlates of attention during the study phase of the experiment. In Experiment X, the evoked potential of the EEG was monitored. In Experiment XI, pupil dilation was used as a measure of mental effort. Both studies were abandoned when it became obvious that they were failures. Whether this was due to the incorrectness of the basic notions behind the experiments or to unreliability of those measures in the particular type of memory situation studied here is not known.

Experiment XII was an attempt to control voluntary attention by signaling the incentive value of a to-be-remembered stimulus with the presence or absence of a tone. The control of attention given to P_2 was effective, but it had no influence on the magnitude of the spacing effect. While more work on this particular point is needed, the outcome suggests that the spacing effect is not produced by a voluntary control process. A review of the literature (Hintzman, 1974) supports this conclusion by revealing the ubiquity of the spacing effect. The effect is found under such a wide variety of conditions--in several different memory tasks, using several dependent variables, and with a variety of to-be-remembered materials--that the possibility that it could be produced by a voluntarily-chosen tactic used by the subject seems rather unlikely.

Experiments XIII and XIV sought evidence for an involuntary attention mechanism--referred to here as habituation and recovery. Both experiments were aimed at manipulating the degree of habituation taking place prior to the spacing interval, to determine whether the form of the spacing (recovery) function was affected. The answer in both cases was quite clear. The form of the spacing function is not affected by either the duration of the presentation prior to the spacing interval or the number of massed repetitions occurring prior to the interval. The spacing effect is apparently only a function of offset to onset time. If a habituation-like process is involved, the degree of habituation apparently cannot be manipulated in this way.

Experiment XV was an attempt to determine whether the spacing effect is due to weaker traces of P_2 being stored at short spacings or to a lower probability of the all-or-none storage of a trace of P_2 . The experiment, in which subjects rated their confidence in their frequency judgments, produced data expected from the incremental hypothesis. However, it was decided after the fact that the data could also be accommodated by a revised version of the all-or-none hypothesis. This attempt to discriminate between incremental and all-or-none processes apparently suffered the same fate as others. Perhaps the issue is not really a testable one after all.

Finally, Experiments XVI and XVI, done by Richard A. Block, in his doctoral dissertation, explored the relationship between memory for duration of an interval and the amount of information retrievable from memory concerning that interval. Neither the number of words that can be recalled from an interval nor the complexity of the sequence of events occurring in

the interval was found to consistently predict memory for duration in both experiments. The results indicate that memory for duration is not related to standard measures in a simple way.

The most important conclusion arising from this research, however, has to do not with specific phenomena or their correct explanations, but with methodology. The primary purpose of this project was to further develop the method of memory judgments and extend the method to the investigation of a wide range of problems. The work undertaken here has shown that memory judgments are a flexible tool that can be adapted to the study of a variety of questions--not only by providing a way of determining what is stored, but also by providing measures that are not subject to some of the limitations (e.g., ceiling effects) characteristic of the more traditional recall and recognition measures. In this regard, then--quite apart from the fate of any particular hypothesis under consideration here--the present project can be judged an unqualified success.

Bibliography

- Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. Psychological Review, 1972, 79, 97-123.
- Anisfeld, M., & Knapp, M. Association, synonymity, and directionality in false recognition. Journal of Experimental Psychology, 1968, 77, 171-179.
- Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.) The Psychology of Learning and Motivation: Advances in Research and Theory. New York: Academic Press, 1968. Vol. 2, pp. 89-195.
- Bernbach, H. A multiple-copy model for post-perceptual memory. In D. A. Norman (Ed.) Models of Human Memory. New York: Academic Press, 1970.
- Bjork, R. A., & Allen, T. W. The spacing effect: Consolidation or differential encoding? Journal of Verbal Learning and Verbal Behavior, 1970, 9, 567-572.
- Block, R. A. Memory and the experience of duration in retrospect. Memory and Cognition, (in press).
- Bray, N. W., & Batchelder, W. H. Effects of instructions and retention interval on memory of presentation mode. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 367-374.
- Conrad, R. Acoustic confusions in immediate memory. British Journal of Psychology, 1964, 55, 75-84.
- Cooper, E. H., & Pantle, A. J. The total-time hypothesis in verbal learning. Psychological Bulletin, 1967, 74, 221-234.
- Glanzer, M., & Cunitz, A. R. Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 351-360.
- Goldwater, B. C. Psychological significance of pupillary movements. Psychological Bulletin, 1972, 77, 340-355.
- Greeno, J. G. Paired-associate learning with short-term retention: Mathematical analysis and data regarding identification of parameters. Journal of Mathematical Psychology, 1967, 4, 430-472.
- Hintzman, D. L. Theoretical implications of the spacing effect. In R. L. Solso (Ed.) Theories of cognitive psychology: The Loyola Symposium. Washington, D.C.: V. H. Winston & Sons, 1974 (in press).

- Hintzman, D. L., & Block, R. A. Repetition and memory: Evidence for a multiple-trace hypothesis. Journal of Experimental Psychology, 1971, 88, 297-306.
- Hintzman, D. L., & Block, R. A. Memory for the spacing of repetitions. Journal of Experimental Psychology, 1973, 99, 70-74.
- Hintzman, D. L., Block, R. A., & Inskip, N. R. Memory for mode of input. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 741-749.
- Hintzman, D. L., Block, R. A., & Summers, J. J. Contextual associations and memory for serial position. Journal of Experimental Psychology, 1973, 97, 220-229.
- Hintzman, D. L., Block, R. A., & Summers, J. J. Modality tags and memory for repetitions: Locus of the spacing effect. Journal of Verbal Learning and Verbal Behavior, 1973 (in press).
- Hintzman, D. L., & Rogers, M. K. Spacing effects in picture memory. Memory and Cognition, 1973 (in press).
- Hintzman, D. L., & Summers, J. J. Long-term visual traces of visually presented words. Bulletin of the Psychonomic Society, 1973, 1, 325-327.
- Hudson, R. L., & Dunn, J. E. A major modification of the Bousfield (1966) measure of category clustering. Behavior Research Methods and Instrumentation, 1969, 1, 110-111.
- Melton, A. W. The situation with respect to the spacing of repetitions and memory. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 596-606.
- Ornstein, R. E. On the experience of time. Harmondsworth, Middlesex, England: Penguin, 1969.
- Peterson, L. R. Search and judgment in memory. In B. Kleinmuntz (Ed.) Concepts and the structure of memory. New York: Wiley, 1967, pp. 153-180.
- Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. Retention of visual and name codes of single letters. Journal of Experimental Psychology Monographs, 1969, 79, No. 1, Part 2.
- Postman, L. A pragmatic view of organization theory. In E. Tulving & W. Donaldson (Eds.) Organization of memory. New York: Academic Press, 1972.
- Postman, L., & Phillips, L. W. Short-term temporal changes in free recall. Quarterly Journal of Experimental Psychology, 1965, 17, 132-138.

- Rundus, D. Analysis of rehearsal processes in free recall. Journal of Experimental Psychology, 1971, 89, 63-77.
- Shaffer, W. O., & Shiffrin, R. M. Rehearsal and storage of visual information. Journal of Experimental Psychology, 1972, 92, 292-296.
- Waugh, N. C. On the effective duration of a repeated word. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 587-595.
- Winograd, E. List differentiation as a function of frequency and retention interval. Journal of Experimental Psychology Monograph Supplement, 1968, 76, No. 2, Part 2.
- Yntema, D. B., & Trask, F. Recall as a search process. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 65-74.